



This is the final peer-reviewed version of the following article:

López Rodríguez, Clara Inés, Pamela Faber, Pilar León Araúz, Juan Antonio Prieto Velasco, and Maribel Tercedor. 2010. La Terminología basada en marcos y su aplicación a las Ciencias Ambientales: los proyectos Marcocosta y Ecosistema. *Arena Romanística* 7, no. 10: 52-74.

You can find more articles authored by LexiCon Research Group members at <<http://lexicon.ugr.es>>.

La Terminología basada en marcos y su aplicación a las Ciencias Ambientales: los proyectos MARCOCOSTA y ECOSISTEMA¹

CLARA INÉS LÓPEZ-RODRÍGUEZ

PAMELA FABER

PILAR LEÓN-ARAUZ

JUAN ANTONIO PRIETO-VELASCO

MARIBEL TERCEDOR-SÁNCHEZ

Abstract

The projects MARCOCOSTA and ECOSISTEMA aim to build a knowledge base representing the field of Environmental Science. To this end, knowledge is extracted from a multilingual corpus, contextual information, concordances, dictionaries and terminological databases, multimedia material (images, animations), and consultation with experts of the CEAMA (Andalusian Center for Environmental Studies). In this paper, we describe the theoretical background (Frame-based Terminology), the methodology and results of these projects, and advocate a descriptive, multidimensional and multimodal approach to knowledge representation and terminology management, in order to facilitate professional communication.

¹ Este trabajo se ha elaborado en el marco del proyecto de excelencia *MARCOCOSTA: Marcos de conocimiento multilingüe en la gestión integrada de zonas costeras* (P06-HUM-01489), financiado por la Junta de Andalucía (España) y por el proyecto I+D *ECOSISTEMA: Espacio Único de Sistemas de Información Ontológica y Tesauro sobre el Medio Ambiente* (FFI2008-06080-C03-01/FILO), financiado por el Ministerio de Ciencia e Innovación de España. Ambos proyectos están dirigidos por la Dra. Pamela Faber de la Universidad de Granada.

Keywords: knowledge representation, Frame-based Terminology, corpus, multidimensionality, multimodality.

Resumen

Los proyectos MARCOCOSTA y ECOSISTEMA tienen como principal objetivo estructurar el conocimiento especializado sobre las Ciencias ambientales y construir una base de conocimiento con diferentes recursos terminológicos multilingües. El conocimiento especializado se extrae de un corpus multilingüe, del análisis de contextos lingüísticos y líneas de concordancia, de diccionarios, bases de datos terminológicas y material multimedia (imágenes y animaciones), y de la consulta a expertos del CEAMA (Centro Andaluz de Medio Ambiente). En este artículo describimos las bases teóricas, la metodología y resultados de estos proyectos, y sostenemos que es necesario abordar la representación del conocimiento y la gestión de la terminología desde una perspectiva descriptiva, multidimensional y multimodal con el fin de facilitar la comunicación entre especialistas.

Palabras clave: representación de conocimiento, Terminología basada en Marcos, corpus, multidimensionalidad, multimodalidad.

1. Introducción

El importante papel que juegan la Ciencia y la Técnica en nuestra vida diaria ha hecho que la terminología y los lenguajes especializados hayan adquirido un gran protagonismo. La Terminología es una disciplina relativamente joven que surgió de la necesidad de procesar, adquirir y transferir conocimiento experto, al tiempo que hacía posible la comunicación especializada. Desde sus inicios, la Terminología ha intentado explicar cómo se representa el conocimiento especializado, cómo se organizan y describen los conceptos especializados y cómo se comportan las unidades terminológicas desde el punto de vista semántico y sintáctico.

En este cometido ha sido fundamental el desarrollo sin precedentes de medios informáticos que permiten gestionar la terminología² en bases de datos, compartir el conocimiento y transmitir el vocabulario especializado de forma inmediata, lo que

² Se entiende la gestión de la terminología como la documentación, almacenamiento, manejo y presentación del vocabulario especializado (Aüstermühl 2001: 102).

facilita la actualización de diccionarios, glosarios, tesauros, y por ende, la comunicación especializada.

Uno de los desafíos actuales de la Terminología es que, frente a una tradicional visión prescriptiva de la terminología, la práctica terminológica y el uso real de los términos han mostrado que es necesario un enfoque descriptivo que estudie las unidades de significación especializadas desde una perspectiva social, lingüística y cognitiva, que tenga en cuenta el dinamismo de la lengua. En segundo lugar, los nuevos soportes multimedia de la información y la interdisciplinariedad de la Ciencia y la Técnica han generado la necesidad de un formato y un modelado adecuados para la descripción multidimensional y multimodal de los términos. El modelado de recursos léxicos y terminológicos debe incluir recursos de información multimodales: a) recursos lingüísticos (léxicos y textuales); y b) recursos multimedia (no lingüísticos): imágenes, audio, vídeo. Por último, en el caso de la terminología en lenguas romances, más concretamente, en español, hay que considerar la escasez de diccionarios especializados actualizados en lengua española y el estatus del inglés como *lingua franca* de la ciencia y técnica.

En este artículo, describiremos las bases teóricas, la metodología y resultados (plasmados en el Tesoro visual *Ecolexicon*) de dos proyectos de investigación sobre Ciencias Ambientales:³

- MARCOCOSTA: Marcos de conocimiento multilingüe en la gestión integrada de zonas costeras (P06-HUM-01489)
- ECOSISTEMA: Espacio Único de Sistemas de Información Ontológica y Tesauro sobre el Medio Ambiente (FFI2008-06080-C03-01/FILO).

Estos proyectos tienen dos objetivos principales. En primer lugar, estructurar el conocimiento especializado sobre las Ciencias Ambientales y generar diferentes recursos terminológicos multilingües:

- a) Una base de datos terminológica trilingüe español-inglés-alemán, que se sustenta en una base de conocimiento.
- b) Un corpus trilingüe que aporte información terminográfica y conceptual.
- c) Un banco de imágenes (Prieto 2007; 2008; 2009).

³ Estos proyectos surgieron a partir de otro proyecto también dirigido por la Dra. Pamela Faber. Se trata de *PUERTOTERM: Ingeniería de puertos y costas: estructuración del conocimiento y generación de recursos terminológicos* (BFF2003-04720), financiado por el Ministerio de Educación y Ciencia.

El segundo objetivo de los proyectos es identificar la estructura conceptual subyacente a las Ciencias Ambientales: los principales conceptos, las relaciones entre conceptos y su multidimensionalidad, es decir, el hecho de que los conceptos puedan ser clasificados de forma diferente, en función del contexto, que activa una u otra faceta del concepto. Los resultados de MARCOCOSTA y ECOSISTEMA están disponibles desde el enlace <http://manila.ugr.es/visual/>

2. Bases teóricas y metodológicas: la Terminología basada en Marcos

Para llevar a cabo los proyectos mencionados en el apartado anterior, hemos seguido un enfoque teórico-práctico denominado *Terminología basada en Marcos* (Faber et al. 2005, 2006, 2007). Este enfoque cognitivo hunde sus raíces en la Gestión de la terminología basada en el conocimiento (*knowledge based terminology management*), iniciada y aplicada con éxito en el proyecto ONCOTERM (Faber 2002, Faber, López y Tercedor, 2001, Oncoterm 2002, López, Faber y Tercedor 2007). La Terminología basada en Marcos (en adelante, TbM) comparte muchos de los presupuestos teóricos de la Teoría Comunicativa de la Terminología (Cabré 1999; 2003) y de la Terminología Sociocognitiva (Temmerman 1997, 2000).

Con la Teoría Comunicativa de la Terminología, la TbM comparte la idea de que hay diferentes “puertas” para acceder a las unidades terminológicas, que podrían compararse con poliedros cuyas caras representan diferentes dimensiones: la cognitiva, la lingüística y la comunicativa. Por otra parte, al igual que la Terminología Sociocognitiva, la TbM acepta la Teoría de prototipos (aunque es demasiado imprecisa para la descripción de las unidades terminológicas) y la estrecha vinculación que hay entre la lengua y los conceptos, al tiempo que reconoce que fenómenos como la polisemia o la sinonimia están presentes en el lenguaje especializado. También comparte con la Terminología Sociocognitiva la idea de que los conceptos y términos evolucionan con el tiempo, y la necesidad de usar ontologías en el proceso de creación de recursos terminológicos.⁴

Por su parte, la Terminología basada en Marcos tiene como objeto de estudio las unidades léxicas y terminológicas tal y como se utilizan en contexto, lo cual requiere conocer el dominio de especialidad y sus principales conceptos, las relaciones

⁴ La base de datos terminológica del proyecto ONCOTERM estaba vinculada a una ontología mediante el programa de gestión y edición de ontologías *OntoTerm*.

proposicionales que se dan en el texto, así como las relaciones conceptuales entre los conceptos del dominio. Para lograrlo se ayuda de la metodología empírica de la Lingüística de corpus.

2.1. Antecedentes de la Terminología basada en Marcos

La TbM surge a partir de modelos lingüísticos y metodológicos sólidos: el Modelo Lexemático Funcional de Martín Mingorance, la Semántica de Marcos de Fillmore y la Lingüística de corpus.

El Modelo Lexemático Funcional o Gramática léxica de Martín Mingorance (1984, 1989, 1995; Faber y Mairal 1999; Faber, López y Tercedor 2001) es una teoría de base léxica que facilita la representación de relaciones conceptuales y colocacionales en el lenguaje general y especializado. Este modelo ofrece una metodología para la organización de conceptos, mediante la información extraída de definiciones lexicográficas y terminográficas estructuradas sistemáticamente. El MLF se ha aplicado con éxito al estudio de la lengua general y especializada, y representa tanto el eje paradigmático como el sintagmático del lenguaje. Al igual que el MLF, la Terminología basada en Marcos considera que las definiciones terminográficas son esenciales a la hora de adquirir conocimiento y ubicar los conceptos dentro de categorías conceptuales más generales, dado que dichas definiciones permiten hacer explícita la pertenencia del concepto a una categoría conceptual, reflejan las relaciones con otros conceptos de la misma categoría, y por último, pueden especificar los atributos y características esenciales de un concepto. Ambos modelos, a la hora de confeccionar definiciones, aplican el análisis definicional y elaboran jerarquías terminográficas. La metodología de establecer jerarquías mediante el análisis de definiciones lexicográficas parte de la premisa básica de que la información contenida en diccionarios y obras terminográficas constituye una red léxico-conceptual que necesariamente tiene correspondencia con el conocimiento expresado. Las definiciones terminográficas se conciben como el puente entre conceptos y términos.

La TbM también se alimenta de la Semántica de Marcos (Fillmore 1982, Fillmore, Johnson y Petruck 2003), aplicada con éxito a proyectos como *Framenet*, que cuenta con una versión española. Fillmore, Johnson y Petruck (2003) definen el marco como una representación esquemática o la representación de un sistema de conceptos relacionados de tal manera que la utilización de un único concepto evoca o activa todo el

sistema conceptual. Y así, en nuestro proyecto, cualquier concepto, por muy especializado que sea, debe encuadrarse en un marco conceptual de referencia que incluye las principales entidades, acciones y relaciones de ese campo de especialidad. A este marco conceptual de referencia se ha denominado EVENTO MEDIOAMBIENTAL, elaborado por Faber, Márquez y Vega (2005). Este evento ha permitido organizar el conocimiento especializado en una estructura basada en marcos, así como la identificación de marcos subordinados más detallados pertenecientes a dominios conceptuales más restringidos semánticamente. El dominio medioambiental tiene un carácter dinámico, dado que intervienen en él una serie de procesos interrelacionados. El EVENTO MEDIOAMBIENTAL describe estos PROCESOS iniciados por un AGENTE (natural o humano) que afecta a otra entidad con función de PACIENTE (las entidades que conforman la costa y superficie terrestre) y produce un RESULTADO. Este fenómeno se puede describir (DESCRIPCIÓN) en torno a una serie de parámetros. Las macro-categorías conceptuales (y algunas otras como INSTRUMENTO, MEDIDA, DISCIPLINA, etc.) y las relaciones conceptuales entre ellas sirven para articular el resto de conceptos de este dominio.

A partir de este modelo, cualquier concepto medioambiental podría encajar en este esquema. Por ejemplo, un AGENTE (humano o natural [como las *olas*]) CAUSAN un PROCESO (natural o artificial, por ejemplo, la *erosión*), que AFECTA a una entidad geográfica con el rol semántico de PACIENTE (tal y como un *acantilado*) y produce un RESULTADO (*playa*). Para describir esos procesos, existe una categoría DESCRIPCIÓN.

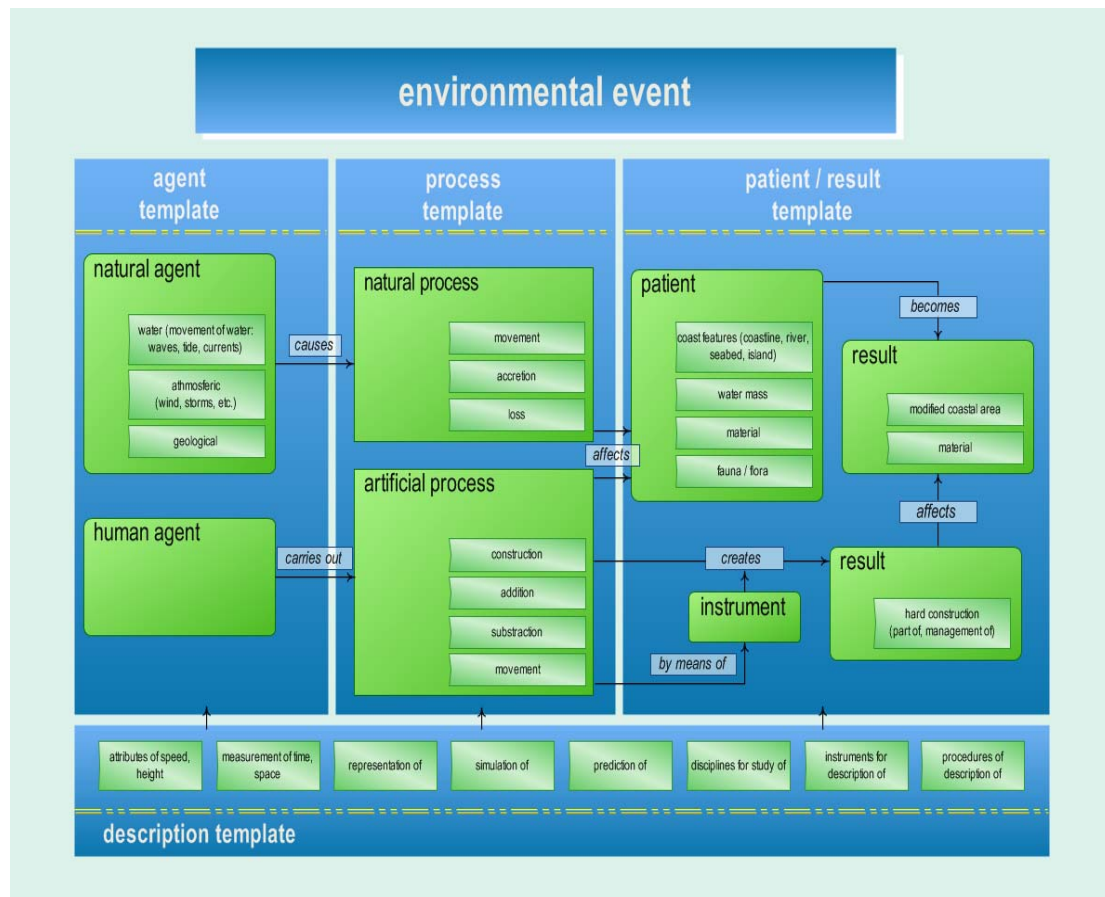


Figura 1. Representación conceptual del EVENTO MEDIOAMBIENTAL (Faber et al. 2006).

Por otra parte, desde finales de la década de 1990, los investigadores de estos proyectos y otros proyectos previos como ONCOTERM y PUERTO TERM han seguido la metodología de la Lingüística de Corpus, es decir, han compilado una ingente cantidad de textos en soporte electrónico y los han analizado con programas de análisis léxico para obtener, entre otros datos empíricos, líneas de concordancia, listas de frecuencia y palabras clave. Con estos programas se facilita la identificación de términos, colocaciones, unidades fraseológicas y patrones sintácticos, se detecta el significado colocacional y prosodia semántica de los términos y se extrae conocimiento sobre el campo de especialidad (Tercedor y López 2008).

3. Metodología: planteamientos generales y fases del proyecto

La representación de conocimiento y la gestión terminológica debe llevarse a cabo combinando una metodología *top-down* —por la que se elabora una lista intuitiva de dominios y subdominios conceptuales *ad hoc* a la que, posteriormente, se le asignarán unidades léxicas o terminológicas— con una *bottom-up*, en la que el

traductor/terminólogo extrae los términos del texto original y los encaja en la red conceptual o evento prototípico representativo del campo de especialidad en cuestión. Según Faber, Tercedor y López (2001), el conocimiento especializado es una subdivisión de nuestro sistema de conocimiento general y, por tanto, se compone de esquemas cognitivos flexibles que permiten la manipulación de ideas y la construcción de hipótesis. Cualquier concepto especializado debe poder encajar en un sistema formado por *eventos* (actividades, acciones, sucesos o situaciones), *objetos* (entidades que no son ni acciones ni propiedades) y *relaciones* que los vinculan entre ellos. Por ejemplo, el concepto AGUA es un objeto, y está vinculado a SALINIDAD mediante la relación ATRIBUTO DE.

Otro de los fundamentos de la metodología es el hecho de que se conciben las imágenes y las definiciones como representaciones del conocimiento. Las definiciones son el núcleo de los proyectos MARCOCOSTA y ECOSISTEMA y siguen un lenguaje claro y conciso. Asimismo, en la elaboración de definiciones y las entradas terminográficas se busca una coherencia *interna* (se unifican las categorías de datos que aparecerán en cada entrada) y *externa* (cómo se interrelacionan entre sí para conseguir un todo).

Si para comprender un campo de especialidad hay que detectar las entidades y acciones básicas de ese campo, que se interrelacionan mediante distintos tipos de relaciones conceptuales hasta construir un sistema conceptual, en MARCOCOSTA y ECOSISTEMA entendemos que dicho sistema conceptual debe quedar reflejado en la definición terminográfica de cada unidad terminológica. La estructuración conceptual del dominio de las Ciencias Ambientales se consigue mediante la elaboración de jerarquías terminográficas, basadas en la extracción de información conceptual de un corpus de textos especializados, de la consulta a recursos terminográficos y a expertos —en concreto, del CEAMA (Centro Andaluz del Medio Ambiente)— y del análisis de material multimedia con imágenes y animaciones. El corpus de textos especializados se analiza con la herramienta de análisis léxico *Wordsmith Tools*,⁵ principalmente con las aplicaciones *WordList*, *Concord*, *KeyWords* y *Viewer*. El análisis del corpus con *Wordsmith Tools* constituye sólo una de las fases de la metodología empleada en los proyectos, que se describen muy brevemente a continuación.

3.1. Recopilación de corpus en soporte electrónico

⁵ Programa creado por Mike Scott [<http://www.lexically.net/wordsmith/index.html>] de la *University of Liverpool* y distribuido por Oxford University Press.

El corpus de nuestro proyecto está formado por textos comparables en inglés, español y alemán que computan aproximadamente 10 millones de palabras. Aunque no disponemos de datos para el subcorpus de alemán, sí podemos afirmar que el corpus en inglés consta de más de 4,5 millones de palabras y el corpus en español, más de 5 millones de palabras. Se trata de un corpus monitor porque nuestra intención es que el corpus siga creciendo.

3.2. Lista de conceptos y términos a partir de la consulta a expertos y del análisis de corpus, recursos terminográficos y material audiovisual

A partir de recursos terminográficos encontrados en Internet sobre ingeniería hidráulica e ingeniería de puertos y costas, la consulta a investigadores del grupo Puertos y Costas de la Universidad de Granada, vinculados al CEAMA, y las listas de frecuencia extraídas con *WordSmith Tools* de nuestro corpus, elaboramos una lista preliminar de los conceptos más importantes en este dominio.

Luego, con la ayuda de recursos multimedia (vídeos, animaciones, imágenes...) y de algunos manuales en soporte papel, fuimos profundizando en ese campo de especialidad mediante la lectura de numerosos textos. En un principio, consultamos textos didácticos tales como manuales o cursos introductorios para posteriormente abordar textos más especializados. Con estas lecturas se fue elaborando un borrador del marco conceptual subyacente al dominio.

3.3. Elaboración de un marco conceptual de referencia mediante la elaboración de jerarquías terminográficas

Como ya hemos dicho en el apartado 2.1, todos términos de las Ciencias Ambientales pueden integrarse en un marco conceptual general, el EVENTO MEDIOAMBIENTAL (Faber et al. 2005). Para llegar a este marco conceptual de referencia, se analizaron las definiciones terminográficas de los conceptos clave del campo de especialidad, con el objetivo de identificar los conceptos y relaciones conceptuales más simples sobre los que se sustenta el sistema conceptual, dado que el ser humano construye el conocimiento especializado a partir de procesos y objetos que le resultan familiares. Los conceptos más simples que se identificaron quedan representados en la Figura 2.

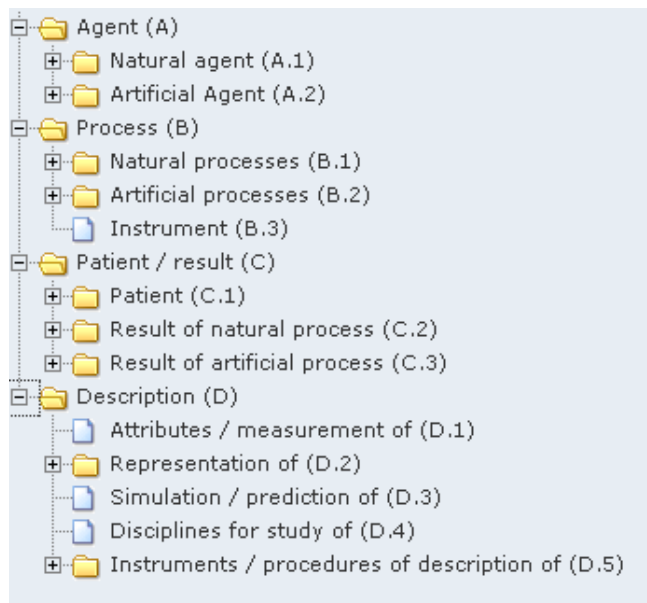


Figura 2. Conceptos básicos del EVENTO MEDIOAMBIENTAL

Además de estos conceptos principales, se identificaron relaciones conceptuales prototípicas como TIPO_DE, RESULTADO_DE, AFECTA_A, las cuales proporcionan coherencia interna a la base de datos, al tiempo que pueden estar verbalizadas en la definición terminográfica. Y así, al definir el concepto *DEPURACIÓN* como un proceso artificial de sustracción (apartado B.2. de la Figura 2), se focalizan las relaciones del Ejemplo 1 y se activan conceptos relacionados, algo que quedaría representado en la base de datos tal y como se indica en la Figura 3.

(1) *DEPURACIÓN*

TIPO_DE: proceso artificial

AFECTA_A: aguas residuales

INSTRUMENTO: depuradora

TIPO_DE *depuradora*: reactor mezclado, reactor de tipo mezclado

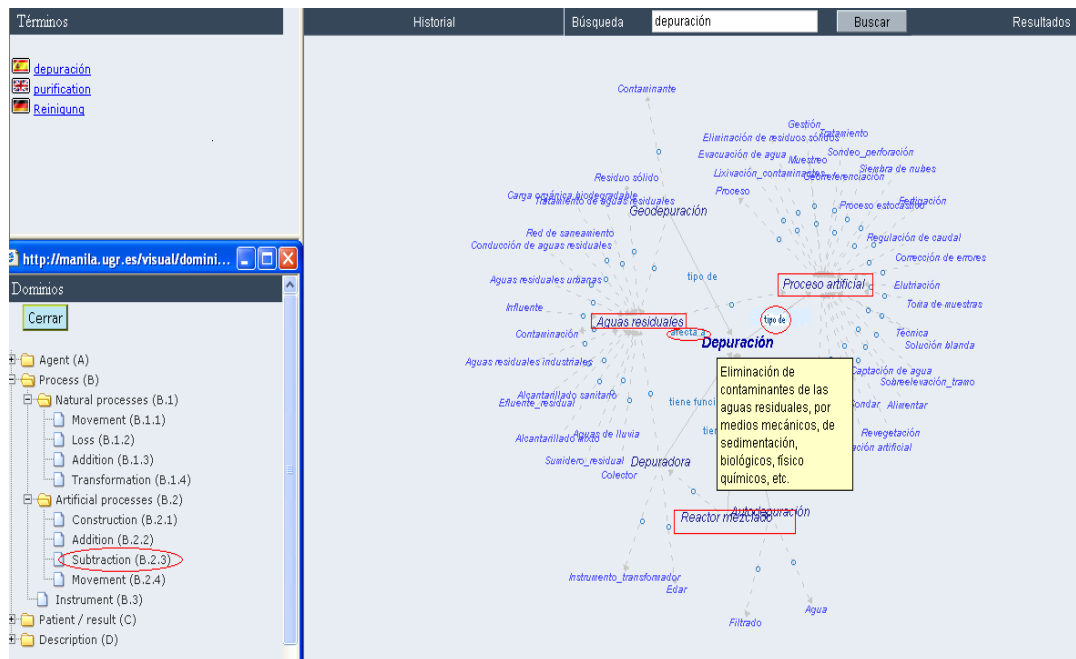


Figura 3. Representación del concepto DEPURACIÓN en *Ecolexicon* (Reimerink y Faber 2009)

3.4. Elaboración de entradas terminográficas y de definiciones coherentes e informativas

En la elaboración de entradas terminológicas se siguen los siguientes pasos:

- inserción en el marco conceptual;
- elaboración de definiciones coherentes e informativas a partir de la aplicación de esquemas categoriales;
- información gramatical básica;
- inserción de ejemplos (contextos) y líneas de concordancia representativas;
- banco de recursos multimedia.

En primer lugar, cada concepto nuevo se ubica en el marco conceptual del EVENTO MEDIOAMBIENTAL y se empiezan a establecer relaciones con otros conceptos. Luego se elaboran definiciones consensuadas de cada lexema mediante el análisis de corpus y la extracción de la información de imágenes y de las entradas de varios diccionarios. La definición terminográfica tiene que presentar de forma explícita el lugar que ocupa el término en la jerarquía, así como representar la estructura conceptual subyacente que comparte un término y sus hipónimos, cohipónimos e hiperónimos.

Para elaborar una red semántica representativa del significado conceptual subyacente en el dominio de la ingeniería de costas, hay que analizar y clasificar

onomasiológicamente sus términos en inglés y español en campos léxicos jerarquizados según las premisas del Modelo Lexemático Funcional. En el MLF, se establece un sistema de definiciones basado en el *Decomposition Principle* de Mel'cuk (1988), que estipula que la definición de una unidad léxica debe contener únicamente términos que sean más simples que dicha unidad léxica.

Por ejemplo, si quisiéramos definir el término *espigón*, tendríamos que consultar las definiciones de diferentes diccionarios especializados y elaborar líneas de concordancia conceptuales que mostraran los parámetros en torno a los cuales se organiza el concepto ESPIGÓN (véase ejemplo 2). A partir de ahí se elaboraría una jerarquía en la que cada concepto se define mediante uno semánticamente más simple; este a su vez, se reduce a uno más simple y así respectivamente, como se ve a continuación:

(2) [ESPIGÓN] TIPO_DE [OBRA DE DEFENSA COSTERA] TIPO_DE [DEFENSA COSTERA] TIPO_DE [OBRA DE INGENIERÍA] TIPO_DE [CONSTRUCCIÓN] TIPO_DE [EVENTO] TIPO_DE [TODO].

Por tanto, términos como *espigón* o *defensa costera* están en una relación de subordinación con *construcción* o *evento*. Esto facilita la ubicación del campo especializado dentro de un contexto más amplio.

Para que las definiciones tengan coherencia interna y externa, se extraen esquemas categoriales o *templates* (Faber, López y Tercedor 2001)⁶ que se aplicarán en la elaboración de las definiciones terminográficas. La ventaja de aplicar esquemas categoriales reside en que estos proporcionan una estructura uniforme a la definiciones, complementan la información codificada en el sistema conceptual y reflejan el sistema conceptual subyacente, como se verá a continuación.

La DEFENSA DE COSTAS es un subdominio importante dentro de la INGENIERÍA DE COSTAS. Se divide en dos categorías: OBRAS DE DEFENSA COSTERA (por ejemplo, *espigón*, *dique* o *muro pantalla*, categorizadas como OBJETOS) y ACCIONES BLANDAS DE DEFENSA COSTERA (por ejemplo, la *regeneración de playas* o la *restauración de dunas*, conceptualizadas como PROCESO). Cada una de ellas presenta una estructura definicional distinta en la que se

⁶ Estos esquemas categoriales tienen bastantes puntos en común con las así llamadas *double-decker definitions* de Fillmore (2003).

basan todos sus subordinados, consiguiendo así un alto grado de coherencia en los distintos niveles jerárquicos del recurso.

En el caso de los OBJETOS, la herencia de propiedades parte de dos relaciones jerárquicas [TIPO_DE y COMPUESTO_DE (MATERIAL)] y dos no jerárquicas [TIENE_FUNCIÓN y UBICADO_EN]; y en el caso de los PROCESOS, la estructura conceptual subyacente a cada definición de sus subtipos consta de dos relaciones jerárquicas (TIPO_DE y FASE_DE, meronimia aplicada a procesos) y tres no jerárquicas (AFECTA_A, TIENE_FUNCIÓN y TIENE_LUGAR_EN), como se plasma en las Figura 4 y 5.

OBRAS DE DEFENSA COSTERA	
	<ul style="list-style-type: none"> ▪ ____ [TIPO_DE] ▪ ____ [UBICADO_EN] ▪ ____ [COMPUESTO_DE (MATERIAL)] ▪ ____ [TIENE_FUNCIÓN]

Figura 4. Esquema definicional de OBRAS DE DEFENSA COSTERA

ACCIONES BLANDAS DE DEFENSA COSTERA	
	<ul style="list-style-type: none"> ▪ ____ [TIPO_DE] ▪ ____ [FASE_DE] ▪ ____ [AFECTA_A] ▪ ____ [TIENE_FUNCIÓN] ▪ ____ [TIENE_LUGAR_EN]

Figura 5. Esquema definicional de ACCIONES BLANDAS DE DEFENSA COSTERA

Así, las definiciones de ESPIGÓN y REGENERACIÓN DE PLAYAS podrían segmentarse conforme a lo representado en las Figuras 6 y 7.

ESPIGÓN: obra de defensa perpendicular a la costa, construida a base de madera, hormigón o escollera, que retarda la deriva litoral y el proceso de erosión.	
	<ul style="list-style-type: none"> ▪ obra de defensa costera [TIPO_DE] ▪ perpendicular a la costa [UBICADO_EN] ▪ madera, hormigón, escollera [COMPUESTO_DE (MATERIAL)] ▪ retardar la deriva litoral y el proceso de erosión [TIENE_FUNCIÓN]

Figura 6. Activación del esquema definicional de OBRAS DE DEFENSA COSTERA en ESPIGÓN

REGENERACIÓN DE PLAYAS:

acción de defensa costera que consiste en realimentar una playa vertiendo nuevo material dragado con el objetivo de proteger la costa contra la erosión, las inundaciones y las tormentas.

- acción blanda de defensa costera [TIPO_DE]
- playa [TIENE_LUGAR_EN]
- realimentación, vertido y dragado [TIENE_FASE]
- material dragado [AFECTA_A]
- proteger la costa de la erosión, las inundaciones y las tormentas [TIENE_FUNCIÓN]

Figura 7. Activación del esquema definicional de ACCIONES BLANDAS DE DEFENSA COSTERA en REGENERACIÓN DE PLAYAS

Toda esta información, caracterizada por la coherencia y la sistematicidad, converge en la interfaz de EcoLexicon del modo siguiente (Figura 8).

The image displays the EcoLexicon interface for the term 'Espigón'. The main window shows a hierarchical network of related terms. The 'Términos' sidebar on the left lists 'espigón', 'groin', and 'groynes'. The 'Detalle de término' window on the right provides detailed information for 'Espigón', including its type, part, material, and function. The 'Recursos' sidebar on the left lists 'Groins [qrobav1f.jpg]' and 'Groins [fascio1f.jpg]'. The 'Detalle de recurso' window on the right shows the details for 'Groins [qrobav1f.jpg]', including its title and description. A diagram in the bottom right corner illustrates the concept of 'GROINS' on a beach, showing how they trap sand and affect longshore transport.

Figura 8. Entrada de ESPIGÓN en interfaz de EcoLexicon

En la Figura 8, la entrada de ESPIGÓN está conectada con sus equivalentes en inglés (*groin, groyne*) y español (*espigón*), su estructura conceptual básica e información gráfica y lingüística (categoría gramatical básica). Dentro de la información lingüística relacionada con el término, el usuario puede encontrar concordancias extraídas del corpus que vuelven a reflejar las relaciones más básicas de su esquema categorial, donde los diferentes valores asociados a cada relación amplían la información contenida en la definición.

En este apartado no desarrollamos exhaustivamente el uso de líneas de concordancia para la extracción de conocimiento sobre la Ingeniería de costas y el medio ambiente, porque este tema ya ha sido abordado en numerosas publicaciones (Faber *et al.* 2006; Tercedor y López 2008).

3.5. Mapas conceptuales limitados al entorno de la entrada terminológica

En la última fase se generan mapas conceptuales limitados al entorno de entradas terminológicas relevantes dentro del EVENTO MEDIOAMBIENTAL. En la Figura 9, se ilustra cómo es posible elaborar submarcos subordinados EVENTO MEDIOAMBIENTAL (Prieto 2007; Prieto 2008). En concreto, se desarrolla el submarco OLEAJE, una estructura que constituye en sí misma un marco dentro de una estructura conceptual superordinada más amplia (el EVENTO MEDIOAMBIENTAL) de la que hereda su configuración de evento y sus macrocategorías (AGENTE, PROCESO, PACIENTE/RESULTADO, DESCRIPCIÓN).

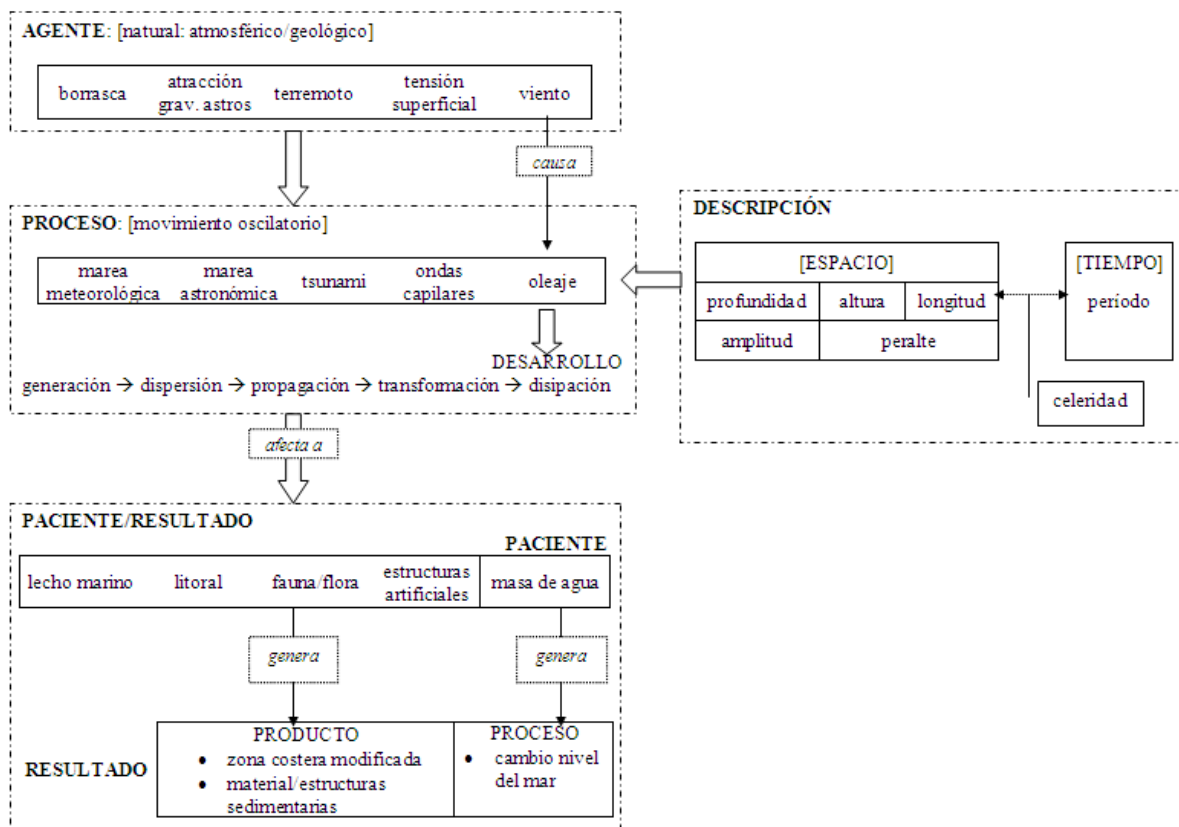


Figura 9. Representación conceptual del marco oleaje (Prieto 2007; 2008)

4. Resultados de los proyectos MARCOCOSTA y ECOSISTEMA

En este apartado mostramos muy brevemente cómo se organizan y describen los conceptos especializados y cómo una base de conocimiento debe aportar un repertorio de información conceptual, lingüística, pragmática y visual.

Además del corpus de textos electrónicos en español, inglés y alemán, el principal resultado del proyecto es un tesoro visual del medio ambiente trilingüe denominado *Ecolexicon*. Se trata de una base de conocimiento de recursos multimedia para la Ingeniería de costas que en la actualidad cuenta con 2838 conceptos y 8628 términos.

4.1. Multidimensionalidad y dinamismo

En *Ecolexicon* plasmamos la multidimensionalidad de los términos de una forma visual y dinámica. La multidimensionalidad se produce cuando un concepto puede clasificarse de manera distinta (o según varias dimensiones) dentro de un mismo sistema conceptual, dependiendo de las características conceptuales utilizadas como base de la subclasificación (Bowker y Meyer, 1993: 123) o, dicho de otro modo, las múltiples

posibilidades de organizar los diferentes conceptos que representan la misma realidad física dentro de un sistema único (Rogers, 2004: 218).

Normalmente, este fenómeno se percibe como una ventaja a la hora de enriquecer los tradicionales sistemas de representación estática, sin embargo, también puede ser la causa de una gran *sobreinformación* en los recursos terminológicos. Éste es el caso de EcoLexicon, donde la orientación al proceso del dominio causa un gran dinamismo y, por tanto, múltiples facetas por representar.

No obstante, el problema de la sobreinformación y la forma de atacarlo responden a una misma cuestión: la incompatibilidad de facetas. Cruse (1995) trata la incompatibilidad de facetas con respecto a la hiponimia. Según uno de sus ejemplos, NOVEL, BIOGRAPHY y TEXTBOOK son hipónimos de BOOK, pero también lo son PAPERBACK y HARDBACK. Los cohipónimos o coordinados de cada grupo se revelan como incompatibles entre sí (si es una NOVEL, no podrá ser un TEXTBOOK), pero no en el caso de comparar dos elementos de ambos grupos (una NOVEL puede ser en PAPERBACK). En este caso, Cruse concluye que las distintas facetas de BOOK se comportan como entidades relacionales distintas (2002: 4).

En este sentido, en *EcoLexicon*, extrapolamos la incompatibilidad de facetas a la totalidad de las relaciones establecidas entre ciertos conceptos según la estructura de sus esquemas categoriales y el contexto. Un ejemplo ilustrativo sería el que proporciona SEDIMENTO. En su conceptualización global, SEDIMENTO causa una sobreinformación imposible de gestionar por el usuario:

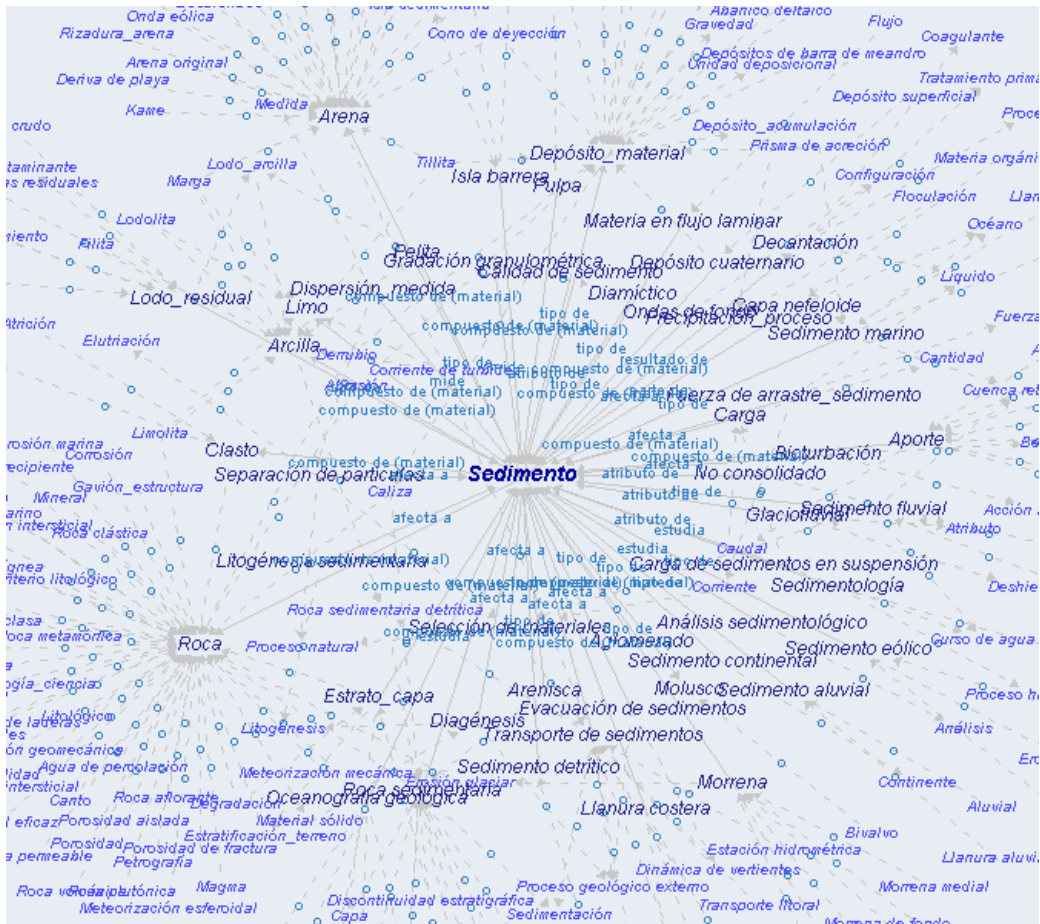


Figura 10. Sobreinformación en torno al concepto SEDIMENTO.

En la Figura 10, SEDIMENTO está vinculado en la misma medida a conceptos relacionados con la geología (MORRENA), con la construcción (GAVIÓN) o con el tratamiento de aguas (DECANTACIÓN). Sin embargo, SEDIMENTO nunca activará muchas de estas proposiciones al mismo tiempo, puesto que en el mundo real evocan situaciones completamente diferentes e incluso incompatibles en términos espacio-temporales (Faber y León Araúz, en prensa).

Esto se explica porque el dominio del Medio Ambiente, de reciente creación como ciencia y de un auge exponencial en los últimos años, aún a numerosos enfoques que abarcan distintas disciplinas, como la ecología, la geología, la química, etc. Así, no es extraño encontrar cierto solapamiento conceptual de origen multidisciplinar, lo que no implica en cualquier caso que las nociones compartidas tengan que ser siempre descritas en los mismos términos.

La aplicación de los marcos a este nivel de representación se basa en la descripción de nueve dominios contextuales (HIDROLOGÍA, GEOLOGÍA, METEOROLOGÍA, BIOLOGÍA, QUÍMICA, INGENIERÍA, TRATAMIENTO DE AGUAS, PROCESOS COSTEROS y NAVEGACIÓN) con

el objetivo de restringir el comportamiento relacional de cada concepto multidimensional.

De este modo, cuando se activan las restricciones, el concepto pasa de mostrar todas sus relaciones potenciales a activar tan sólo aquellas compatibles según el contexto. En la Figura 11, SEDIMENTO está reconceptualizado según el dominio TRATAMIENTO DE AGUAS.

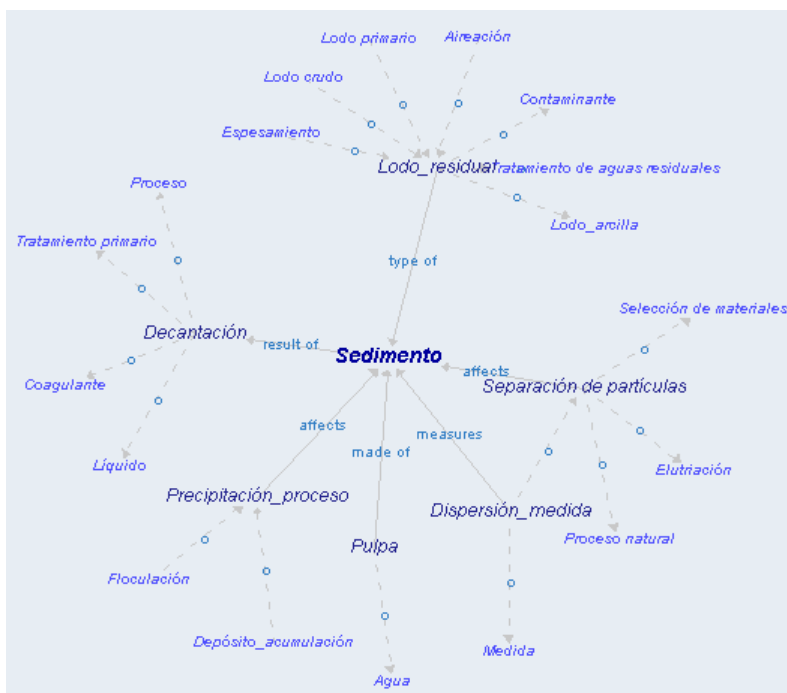


Figura 11. Reconceptualización de SEDIMENTO en el contexto del TRATAMIENTO DE AGUAS.

4.2. Multimodalidad

En nuestro proyecto incluimos imágenes como mecanismo visual de representación del conocimiento en recursos de gestión terminológica, dado que permiten evocar y activar la estructura conceptual del dominio de forma más rápida y complementaria a la información semántica, sintáctica y pragmática codificada textualmente en cada entrada terminológica (Prieto 2008).

De acuerdo con el W3C, la multimodalidad es un proceso por medio del cual los usuarios pueden interactuar con la red y acceder por completo a la información (W3C 2005). Los recursos terminográficos multimodales como *EcoLexicon* proporcionan nuevas vías de recepción de la información y comprensión de los conceptos. Imágenes, vídeos, sonidos, etc. forman parte de nuestro contacto diario con los medios y han propiciado que los usuarios de los lenguajes de especialidad adapten las herramientas con las que contribuyen al intercambio de información especializada.

Por ello, proporcionar información en distintas lenguas, en diversos formatos y desde diferentes perspectivas facilita la comunicación profesional y hace accesible el conocimiento especializado al público general. En MARCOCOSTA Y ECOSISTEMA se representa la multimodalidad en función de la relación entre los recursos visuales y la información lingüística, conceptual y contextual. Por ello, las imágenes deberían:

- estar vinculadas de manera coherente con otras categorías de datos en las que se incluye información multimodal;
- centrarse en los aspectos más significativos de las definiciones terminográficas;
- destacar las relaciones conceptuales más relevantes.

La relación entre la información gráfica y la información lingüística de las definiciones se hace evidente en la representación visual de las relaciones conceptuales que evocan el rol semántico del concepto, así como en la representación de conceptos con diferentes grados de especificidad. La Figura 12 constituye un claro ejemplo de la convergencia de información en términos de multimodalidad.



Figura 12. Campaña LUCHEMOS CONTRA LA DESERTIFICACIÓN⁷ (Prieto y López 2009: 195)

De acuerdo con las categorías que conforman el EVENTO MEDIOAMBIENTAL, podemos definir la *desertificación* como un PROCESO [ROL SEMÁNTICO PROTOTÍPICO] por el que los *suelos* [PACIENTE] de *zonas áridas* [UBICACIÓN_LUGAR] se degradan como

⁷ Fuente: Ministerio argentino de Medio Ambiente y Desarrollo Sustentable.

consecuencia de la *acción del hombre* o de *agentes naturales* [AGENTE] dando lugar a *tierras de baja calidad e infértiles* [RESULTADO]. En consecuencia, creemos que los conceptos y los roles semánticos que evoca el término *desertificación* o *desertización* deberían hacerse explícitos, de un modo u otro, tanto en la definición del término como en las imágenes que lo ilustran.

Las imágenes desempeñan una función muy importante a la hora de adaptar el grado de especificidad de los conceptos especializados al conocimiento que los receptores poseen en un determinado campo temático. En consecuencia, podemos definir los conceptos de manera distinta si estos pertenecen a dominios distintos, especialmente si son multidimensionales, pues diferentes puntos de vista ayudan a focalizar las diferentes dimensiones. La representación gráfica de conceptos especializados es un recurso extremadamente útil para explicitar este tipo de información.



Figura 13. Imagen auxiliar para la definición de DESERTIFICACIÓN [neutra] (Prieto y López 2009: 197)

Hay por tanto algunas diferencias entre las imágenes que representan conceptos pertenecientes al nivel general de conceptualización y dirigidas a una audiencia legla (Figura 13), e imágenes que representan conceptos que pertenecen a niveles de conceptualización más específicos y dirigidas a expertos (Figura 14).

[Meteorología] Desertificación: Tendencia de una región a adquirir las características meteorológicas propias de zonas desérticas.	[Ecología; Agricultura] Desertificación: Proceso por cual un terreno fértil deja de serlo al tiempo que se convierte en un desierto.
--	--

⁸ http://www.meted.ucar.edu/hydro/basic/HydrologicCycle/print_version/03-atmospheric_water.htm

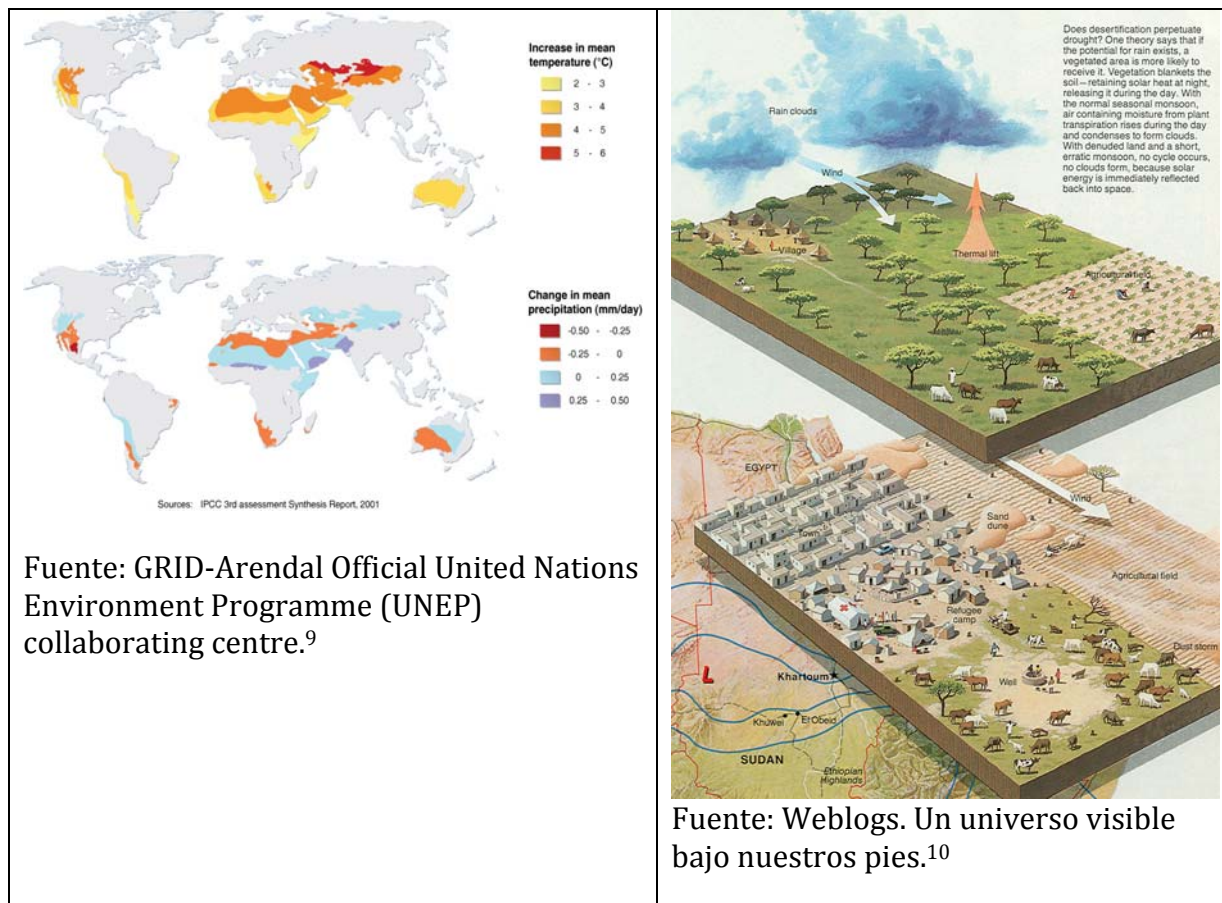


Figura 14. Imagen auxiliar para la definición de DESERTIFICACIÓN [para experto] (Prieto y López 2009: 198)

Estas diferencias parecen apuntar a los aspectos semióticos y cognitivos de la información gráfica, tales como el grado de semejanza entre la imagen y la entidad o proceso representado y la dificultad para comprender el contenido semántico de la imagen. La información lingüística de las definiciones puede ayudar a determinar el nivel de especialización de los textos y, por tanto, a saber cómo adaptar las imágenes de modo que satisfagan las expectativas de los usuarios finales, faciliten el acceso al conocimiento especializado y promuevan la creación de una conciencia medioambiental responsable.

4.3. Variación terminográfica: motivación y representación

Uno de los retos a los que nos enfrentamos al inicio del proyecto fue hacer frente a la ausencia de representación de relaciones de sinonimia en las bases de datos existentes

⁹ http://maps.grida.no/go/graphic/climate_change_scenarios_for_desert_areas

¹⁰ <http://weblogs.madrimasd.org/universo/gallery/image/2642.aspx>

en los dominios objeto de estudio. Puesto que uno de los objetivos de la terminología basada en marcos es entender la representación del conocimiento desde la descripción de cómo se usa la lengua, se hacía necesario (a) determinar qué términos representaban mejor un concepto determinado y (b) recoger las variantes terminológicas de los términos principales en la base de datos. Una variante terminológica es una forma que está semántica y conceptualmente relacionada con el término principal (Daille 1996: 201). Para acceder a las variantes terminológicas, el contexto nos ofrece los matices semánticos y pragmáticos necesarios para decidir su pertinencia. Además, el contexto nos ofrece también información sobre el lugar que ocupa un concepto y las especificaciones de un dominio (Barsalou 2003). Por lo tanto, la motivación para la presencia de variantes terminológicas en el discurso especializado no solo responde a razones pragmáticas, sino también cognitivas, puesto que una variante se puede entender como la manifestación lingüística de la multidimensionalidad conceptual: se da prevalencia a una determinada dimensión del concepto en una determinada denominación. Así, los términos “agua salada” y “agua marina” representan ambos el mismo concepto, pero en el caso del primero se resalta la dimensión de la constitución del agua, mientras que el segundo, es la localización la dimensión prominente. Desde la perspectiva pragmática, también era necesario recoger términos que actúan como variantes dependientes del registro (Tercedor y Méndez 2000) tales como: *cadena trófica* y *cadena alimentaria*. Además, también forman parte de la base de conocimiento usos geográficos tales como *bufadera* (México) frente a *bufadero* (España).

5. Conclusiones

Sustentada sobre un sólido marco teórico, la Terminología basada en Marcos logra representar la multidimensionalidad de los términos con un rico repertorio de información conceptual, lingüística, pragmática y visual. En este artículo hemos descrito las bases teóricas, la metodología y resultados de los proyectos MARCO COSTA y ECOSISTEMA, que han puesto de manifiesto que es necesario abordar la representación del conocimiento y la gestión de la terminología desde una perspectiva descriptiva, multidimensional y multimodal con el fin de facilitar la comunicación entre especialistas.

6. Bibliografía

Austermühl, Frank. 2001. *Electronic tools for translators*. Manchester: Saint Jerome.

- Cabré, M. Teresa. 1999. Terminology. Theory, Methods, Applications. Amsterdam: John Benjamins. [Traducido por Janet Ann DeCesaris].
- Barsalou, Laurence W. 2003. Situated simulation in the human conceptual system. *Language and cognitive processes* 18 5/6, 513-562.
- Bowker, L y Meyer, I. 1993. Beyond Textbook. Concept systems: handling multidimensionality in a new generation of term banks. Schmitz, K.D. (ed.), TKE '93 Terminology and Knowledge Engineering. Frankfurt/a.M: Indeks Verlag. 123-137.
- Daille, Beatrice *et al.* 1996. Empirical observation of term variations and principles for their description. *Terminology* 3 (29). 197-257.
- Cabré, M. Teresa. 2003. Theories of Terminology: their description, prescription and explanation. *Terminology* 9 (2): 163-199.
- Cruse, D.A. 2002. Hyponymy and its varieties. Green, R., Bean, C.A. y Myaeng, S.H. (eds.), *The semantics of relationships*. Dordrecht: Kluwer. 3-21.
- Cruse, D.A. 1995. Polysemy and related phenomena from a cognitive linguistic viewpoint. St. Dizier, P. y Viegas, E. (eds.), *Computational Lexical Semantics*. Cambridge: Cambridge University Press. 33-39.
- Faber, Pamela. 2002. Terminographic definition and concept representation. Maia, Belinda, Haller, Jonathan y Urlrych, Margherita (eds.) *Training the Language Services Provider for the New Millennium*, Porto: Faculdade de Letras, Universidade do Porto, 343-354.
- Faber, Pamela y R. Mairal Usón. 1999. *Constructing a lexicon of English verbs*. Berlin: Mouton de Gruyter.
- Faber, Pamela, Clara López-Rodríguez y Maribel Tercedor-Sánchez. 2001. "La utilización de técnicas de corpus en la representación del conocimiento médico". *Terminology* 7 (2), 167-197.
- Faber, Pamela, C. Márquez-Linares y Miguel Vega-Expósito. 2005. Framing Terminology: A process-oriented approach. *META* 50 (4). <<http://www.erudit.org/livre/meta/2005/000255co.pdf>> 31.1.2010.
- Faber, Pamela *et al.* 2006. Process-oriented terminology management in the domain of Coastal Engineering. Número especial de *Terminology*, vol. 12 (2). "Processing of terms in specialized dictionaries: new models and techniques". 189-213.
- Faber, Pamela *et al.* 2007. Linking images and words: the description of specialized concepts. *International Journal of Lexicography*. Volume 20. 39-65
- Faber, Pamela y León Araúz, Pilar (en prensa). "Dinamismo en las bases de conocimiento terminológico". *Íkala*.
- Fillmore, C. J. 1982. Frame semantics. Linguistic Society of Korea (ed.) *Linguistics in the Morning Calm*. Seoul: Hanshin Publishing Co, 111-137.
- Fillmore, C.J. 2003. Double-decker definitions: The role of frames in meaning explanations. *Sign Language Studies* 3: 3, 263-295.
- Fillmore, C. J., C. R., Johnson y Petruck, M. 2003. Background to FrameNet. *International Journal of Lexicography*. vol. 16 (3). Oxford Journals.
- López Rodríguez, Clara Inés, Pamela Faber Benítez y Maribel Tercedor Sánchez. 2006. Terminología basada en el conocimiento para la traducción y la divulgación médicas: el caso de Oncoterm. *Panacea@*. Vol. VII, n. 24. Diciembre, 2006. <http://www.medtrad.org/panacea/IndiceGeneral/n24_tradyterm-l.rodriguez.etal.pdf>. 31.1.2010.
- Martín Mingorance, L. 1984. "Lexical fields and stepwise lexical decomposition in a contrastive English-Spanish verb valency dictionary". Hartmann (ed.), *LEXeter*

- 83: Proceedings of the International Conference on Lexicography. Tübingen: Niemeyer, 226-236.
- Martín Mingorance, L. 1989. "Functional Grammar and Lexematics". Tomaszczyk, J. and B. Lewandowska (eds.), *Meaning and Lexicography*. Amsterdam / Philadelphia: John Benjamins, 227-253.
- Martín Mingorance, L. 1995. "Lexical logic and structural semantics: methodological underpinnings in the structuring of a lexical database for natural language processing". Hoinkes (eds.), *Panorama der Lexikalischen Semantik*. Tübingen: Gunter Narr, 461-474.
- ONCOTERM 2002. "Sistema Bilingüe de Información y Recursos Oncológicos". <<http://www.ugr.es/~oncoterm/oncodesc.htm>>31.1.2010.
- Mel'cuk, I. 1988. "Semantic description of lexical units in an explanatory combinatorial dictionary: Basic principles and heuristic criteria". *International Journal of Lexicography* 1 (3), 165-188.
- Prieto-Velasco, J. A. 2007. *Hacia un modelo de uso de ilustraciones en el discurso científico-técnico*. *Sendebarr* 18, 127-158.
- Prieto-Velasco, J. A. 2008. *Información gráfica y grados de especialidad en el discurso científico-técnico: un estudio de corpus*. Tesis doctoral. Universidad de Granada. <<http://hera.ugr.es/tesisugr/17491332.pdf>>
- Prieto-Velasco, J. A. 2009. *Traducción e imagen: la información visual en textos especializados*. Ediciones Tragacanto: Granada.
- Prieto-Velasco, Juan Antonio y Clara I. López-Rodríguez. 2009. "Managing graphic information in terminological knowledge bases". *Terminology* 5 (2). Ámsterdam/Filadelfia: John Benjamins. 179-213.
- Reimerink, Arianne y Pamela Faber. 2009. *EcoLexicon: A Frame-Based Knowledge Base for the Environment*. J. Hřebíček et al. (eds.), *Towards eEnvironment. Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe*. Praga: Masaryk University.
- Rogers, M. 2004. "Multidimensionality in concepts systems: a bilingual textual perspective". *Terminology* 10: 2, 215-240.
- Temmerman, Rita. 1997. Questioning the univocity ideal. The difference between sociocognitive Terminology and traditional Terminology. *Hermes, Journal of Linguistics* 18: 51-91.
- Temmerman, Rita. 2000. *Towards new ways of Terminology Description: The sociocognitive approach*. Amsterdam/Philadelphia: John Benjamins.
- Tercedor Maribel y Beatriz Méndez. 2000. *Fraseología y variación terminológica: estudio descriptivo en corpora biomédicos*. *Terminologie et Traduction*. 2, 2000. 82-100.
- Tercedor-Sánchez, Maribel y Clara Inés López-Rodríguez. 2008. "Integrating corpus data in dynamic knowledge bases: The PUERTOTERM Project". *Terminology* 14 (2). 159-182. Ámsterdam/Filadelfia: John Benjamins.